(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 January 2004 (15.01.2004)

PCT

(10) International Publication Number
WO 2004/006559 A2

(51) International Patent Classification[7]: H04N

(21) International Application Number: PCT/US2003/021650

(22) International Filing Date: 9 July 2003 (09.07.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/394,630	9 July 2002 (09.07.2002)	US
60/394,922	9 July 2002 (09.07.2002)	US
60/394,588	9 July 2002 (09.07.2002)	US
10/356,692	31 January 2003 (31.01.2003)	US
10/356,322	31 January 2003 (31.01.2003)	US
10/377,266	28 February 2003 (28.02.2003)	US
10/378,046	28 February 2003 (28.02.2003)	US

(71) Applicant: KALEIDESCAPE, INC. [US/US]; 339 North Bernardo Avenue, Suite 100, Mountain View, CA 94043 (US).

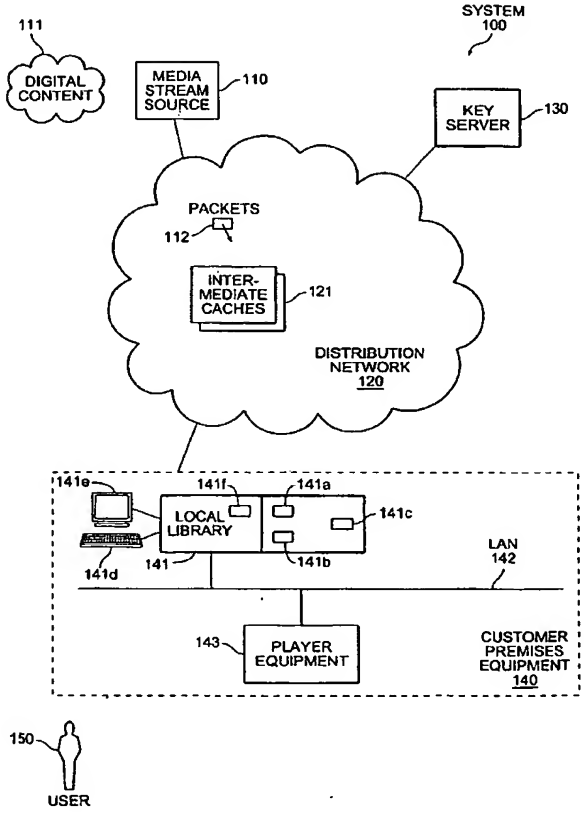
(72) Inventors: MALCOLM, Michael, A.; P.O. Box 7667, Aspen, CO 81612 (US). COLLENS, Daniel, A.; 790 Bonavista Drive, Waterloo, Ontario N2K 3Z8 (CA). WATSON, Stephen; 65 Clinton Street, Toronto, Ontario M6G 2Y4 (CA). RECHSTEINER, Paul; 109 Front St. E., Apt. 627, Toronto, Ontario M5A 4P7 (CA). HUI, Kevin; 308-29 West Avenue, Kitchener, Ontario N2M 5E4 (CA).

(74) Agent: SWERNOFSKY, Steven, A.; Swernofsky Law Group PC, P.O. Box 390013, Moutain View, CA 94039-0013 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: SECURE PRESENTATION OF MEDIA STREAMS IN RESPONSE TO ENCRYPTED DIGITAL CONTENT



(57) Abstract: Secure presentation of media streams includes encoding the media streams into digital content, encrypting a portion of that digital content, the portion being required for presentation, in which the encrypted version is substantially unchanged in formatting parameters from the clear version of the digital content. Selecting those portions for encryption so there is no change in distribution of the media stream: packetization of the digital data, or synchronization of audio with video portions of the media stream. When encoding the media stream into MPEG-2, refraining from encrypting information by which the video block data is described, packet formatting information, and encrypting the video block data using a block-substitution cipher. A block-substitution cipher can be used to encrypt each sequence of 16 bytes of video data in each packet, possibly leaving as many as 15 bytes of video data in each packet in the clear.

WO 2004/006559 A2

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

SECURE PRESENTATION OF MEDIA STREAMS IN RESPONSE TO ENCRYPTED DIGITAL CONTENT

Background of the Invention

1. *Field of the Invention*

The invention relates to presentation of media streams in response to digital content.

2. *Related Art*

Distribution of digital content representing media streams, such as for example movies, is subject to several problems. One problem is that it is easy to make exact copies of digital content, thus allowing any recipient of that content to redistribute it, whether authorized or not. It would be advantageous to be able to distribute digital content, particularly digital content representing media streams, without fear of its unauthorized distribution. This would be particularly advantageous when it is desired to distribute digital content using a communication link, such as for example a computer network or other technique for distribution to end viewers (for example, either on demand, in anticipation of future demand, or in response to something else).

One known solution is to encrypt the digital content that represents the media stream, so that a recipient of that digital content cannot easily redistribute it in a readily presentable (that is, unencrypted) format to unauthorized recipients. However, even when digital content is distributed in an encrypted form, it must be decrypted before it can be presented to a viewer. Thus, there is at least some time for each movie, during distribution from originator to viewer, during which that movie is available in an unencrypted format (herein sometimes also called "in the clear"). At times, and in places in any presentation system, when that movie is available in the clear, that movie is vulnerable to security attacks. For example, an unauthorized person might copy the movie in its unencrypted format and distribute or use it without authorization.

Accordingly, it would be advantageous to provide a method (and devices for performing it) by which the digital content can be used for presentation as a media stream, without exposing that digital content in the clear. However, there are several issues related to achieving this goal.

- It would be desirable for the device to be relatively tamper-resistant, so that the work factor for obtaining the digital content in the clear would be substantially greater than simply purchasing copies (or at least, greater than other possibly available techniques for unauthorized procurement).

- It would also be desirable for the device to expose the digital content representing the media stream as little as possible. For some examples, having the digital content (or a key from which that digital content could be obtained) in the clear in a memory would be less desirable than only having the digital content in the clear on an internal bus, which itself would be less desirable than only having the digital content in the clear when actually presented on a screen for viewing by an end-user.

These issues present a need for separating that part of the device that has access to keys for decryption into a separate set of "trusted" hardware and software elements, with the effect that it would be advantageous for at least some of the device to be implemented in tamper-resistant hardware operating under control of verified software.

- It would be desirable for the device to be able to both decode digital content representing media streams, and to provide common playback functions known for media streams, without these functions involving complete decryption of the digital content. These functions might include navigation within the digital content (such as for example fast-forward and rewind functions), content selection within the digital content (such as for example chapter-skip and multi-angle selection functions), or manipulation of the presentation (such as for example freeze-frame or single-frame-advance functions).

- It would be desirable for the device to be able to provide access to metadata about the one or more media streams, such as a title or rating, or other information about the

media streams for which it is generally acceptable to maintain that information in the clear, without these functions involving complete decryption of the digital content.

- It would be desirable for the device to be able to provide differing access to distinct end-users for selected portions of one or more media streams, such as for example differing access to audio versus video, or English-language versus French-language versions, or US releases versus UK releases, or "airline" versions versus "general release" versions, for the same media stream, without these functions involving complete decryption of the digital content.

It would be desirable for these playback functions, and possibly others, to be implemented in relatively unverified software. In one embodiment, only verified hardware or software would be allowed access to keys for decrypting the digital content. However, there are many such functions for which it would be desirable to have them be available to the user, without having those functions be implemented in tamper-resistant hardware (which would be more expensive, and would be difficult to update), or in verified software (which would also be more difficult to update, and might also be more expensive to create).

Formats now used for encoding digital content representing a media stream for digital distribution (such as for example MPEG-1, MPEG-2, and MPEG-4) are relatively complex. These formats provide for dividing up the digital content into multiple packets. Thus, it is possible when parsing digital content representative of media streams, that encryption might involve maintaining substantial state information across many such packets. A device able to conduct both the parsing and stitching operations might need substantial working memory. In general, having to maintain less state across packet boundaries would allow the hardware and software for decoding and decrypting the encoded and encrypted movie to be simpler, and would allow the digital content for the movie to be less exposed in the clear.

Formats used for encoding digital content representing media streams also provide for partial delivery of portions of the digital content at different times, such as when sending the digital content is interrupted and later restarted, or when packets including portions of the digital content arrive out of order, or with parts missing. Similar to the

problem involving multiple packets, a device able to recover from partial delivery of only a portion of the digital content might need to maintain substantial state, or to maintain substantial working memory. In general, having to maintain less state across packet boundaries would allow the hardware and software for decoding and decrypting the encoded and encrypted movie to be more robust with regard to handling packets that arrive out of order, or with parts missing.

Formats used for encoding digital content representing media streams provide for additional information about the media stream, such as a title, for which it might be advantageous to have available even when the media stream is not actually being presented to the viewer. For example, it might be advantageous to allow a potential viewer to browse titles and related information, or even to conduct a computerized search on that information, without actually presenting the media stream. A device able to provide that information rapidly, such as on a random access basis with regard to the digital content representing that media stream, would involve substantial resources for computation and memory, likely relatively proportionate to the amount of the digital content desired to be reviewed on a random access basis, with the effect that such a device would thus be relatively insecure against attack, as either decryption keys or digital content in the clear would be available to those parts of the system for which such random access were desired.

Accordingly, it would be advantageous to provide an improved technique for presentation of digital content representing a media stream, such as the technique in which devices able to access the digital content are not allowed access to the media stream represented by that digital content, but still are allowed access to metadata regarding that media stream.

Summary of the Invention

A method of secure presentation of media streams in response to encrypted digital content includes (1) encoding the media stream into a digital content format representing that media stream, (2) encrypting a portion of that digital content, less than the entire digital content format representing that media stream, the portion of the digital content that is encrypted being required for presentation of the media stream, (3) in which the

encrypted version of the digital content is substantially unchanged in formatting parameters from the clear version of the digital content.

Formats used for encoding digital content representing media streams provide for encapsulating information in a hierarchy of layers, each relatively higher-level layer representing an abstraction for which each relatively lower-level layer represents an implementation thereof. As described herein, in an aspect of the invention, the highest-level layer (or multiple higher-level layers) represent audio and video information for the media stream, while relatively lower-level layers represent techniques by which that information is broken into packets, indexed, multiplexed, and supplemented with metadata (such as for example closed captioning and copyright information). As described herein, in an aspect of the invention, only the audio and video information for the media stream is encrypted, while other relatively lower-level layers remain "in the clear" (that is, unencrypted).

More generally, formats used for encoding digital content representing media streams provide a tree-structure in which information is disposed, the audio and video data being incorporated into leaves of the tree and various types of metadata (such as for example control information) being incorporated into branches of the tree. After reading this application, those skilled in the art will recognize that a tree structure is not the only possible format, and that in general, any partial ordering of information might be specified by a format used for encoding digital content representing media streams, where the audio or video data are specified to have a selected ordering with regard to metadata for that digital content.

As described herein, in an aspect of the invention, where that format used for encoding the digital content can be represented as a tree, it suffices for a subtree of the digital content closed root-ward to be unencrypted. In this context, "closed root-ward" describes the case where if a node X in the tree T is included in a set of nodes (and thus unencrypted), so is every node in a path from X toward the root of the tree T. In one embodiment, substantially all the leaves of the tree T are encrypted, and the system is still able to parse the MPEG stream, with the only limitation being that the system cannot present the actual audio or video without decryption of those leaves.

Similarly, where that format used for encoding the digital content can be represented as a partial ordering, it suffices for a portion of that partial ordering closed backward under that partial ordering to be unencrypted. In this context, "closed backward" describes the case where if an element X in the partial ordering P is included in a set of elements (and thus unencrypted), so is every element Y for which $Y < X$ in the partial ordering P. In one embodiment, substantially all the audio and video elements of the partial ordering P are encrypted, and the system is still able to parse the MPEG stream, with the only limitation being that the system cannot present the actual audio or video without decryption of those audio and video elements.

More generally, in this context "encrypted" and "unencrypted" can be replaced with distinct levels of hardness to decode the associated elements X and Y without having a presentation device key. For one example, not intended to be limiting in any way, the audio and video elements of the tree T (or the partial ordering P) might be encrypted using the AES-128 block cipher, while the control elements, MPEG packet headers, and MPEG pack headers might be encrypted using a substantially less secure technique, such as a bitwise XOR with a selected password. As described above, so long as the less-strongly encrypted portions form a collection that is closed root-ward (for a tree T) or closed backward (for a partial ordering P), the system will be able to parse the MPEG stream in relatively non-secure hardware and software, while still being limited to a relatively secure portion with the appropriate key to present audio and video.

After reading this application, those skilled in the art will recognize that more generally, "encryption" can be replaced by any security technique, such as for example physical hardware security such as hidden mask layers in a ROM or ASIC. For one example, multiple levels of security might include (a) a first level readable like a file in a computer; (b) a second level readable only by coupling a probe to an external port of the presentation device, (c) a third level readable only by coupling a probe to an internal bus of the presentation device, (d) a fourth level readable only by emulation of the circuitry of the presentation device, and (e) a fifth level readable only by reverse engineering of the integrated circuit and examination with an electron microscope.

An aspect of the method includes selecting those portions of the digital content for encryption so that there is no substantial change in distribution of the digital content representing the media stream, such as for example (1) packetization of the digital data, or (2) synchronization of audio with video portions of the media stream. In a preferred embodiment, unchanged distribution can be accomplished by making no substantial change in length of portions of the video packet data, such as for example individual packets of an MPEG-encoded movie.

In one embodiment, the method includes, when encoding the media stream into a digital content format, such as for example MPEG-2, (1) refraining from encrypting information by which the video packet data is described, such as for example packet formatting information, and (2) encrypting the video packet data using a block-substitution cipher. For example, a block-substitution cipher can be used to encrypt each sequence of 16 bytes of video data in each packet, possibly leaving as many as 15 bytes of video data in each packet in the clear. In one embodiment, the method includes (3) separately encrypting the audio portion of the media streams, and possibly other selected data portions of the media streams, within the digital content, with the effect that these separate data portions of the media streams might be made differently available to distinct selected users or groups of users.

The invention is not restricted to movies, but is also applicable to other media streams, such as for example animation or sound, as well as to still media, such as for example pictures or illustrations, and to databases and other collections of information.

Brief Description of the Drawings

Figure 1 shows a block diagram of a system for secure presentation of media streams in response to encrypted digital content.

Figure 2 shows a process flow diagram of a method for secure presentation of media streams in response to encrypted digital content.

Detailed Description of the Preferred Embodiment

In the description herein, a preferred embodiment of the invention is described, including preferred process steps and data structures. Those skilled in the art would realize, after perusal of this application, that embodiments of the invention might be implemented using a variety of other techniques not specifically described, without undue experimentation or further invention, and that such other techniques would be within the scope and spirit of the invention.

Lexicon

The general meaning of each of these following terms is intended to be illustrative and in no way limiting.

- The phrase "media stream" describes information intended for presentation in a sequence, such as motion pictures including a sequence of frames or fields, or such as audio including a sequence of sounds. As used herein, the phrase "media stream" has a broader meaning than the standard meaning for "streaming media," (of sound and pictures that are transmitted continuously using packets and that start to play before all of the content arrives). Rather, as described herein, there is no particular requirement that "media streams" must be delivered continuously. Also as described herein, media streams can refer to other information for presentation, such as for example animation or sound, as well as to still media, such as for example pictures or illustrations, and also to databases and other collections of information.

- The phrase "digital content" describes data in a digital format, intended to represent media streams or other information for presentation to an end viewer. "Digital content" is distinguished from packaging information, such as for example message header information. For the two phrases "digital content" and "media stream," the former describes a selected encoding of the latter, while the latter describes a result of presenting any encoding thereof.

- The phrase "embedded information in a media stream" describes information incorporated into a set of digital content representing that media stream, in a form capable of later detection. For example, digital content representing media streams might include embedded information, such that the media streams are still capable of presentation to a viewer without substantial change, but in which the embedded information can be recovered by suitable processing of the digital content.

- The phrase "embedding information in a media stream" describes generating a set of digital content representing that media stream, for which the digital content both represents the media stream and also includes the embedded information in a form capable of later detection.

- The term "watermark" describes a schema for digital content by which information can be embedded into that digital content. In preferred embodiments, as described in related applications "Parallel Distribution and Fingerprinting of Digital Content," (Collens et al.) U.S. application number 10/356,692, filed 31 January 2003, and "Watermarking and Fingerprinting Digital Content Using Alternative Blocks to Embed Information," (Watson et al.) U.S. application number 10/356,322, filed 31 January 2003, an attacker cannot easily remove the watermark. However, the concept of a watermark as described herein is sufficiently general to include watermarks that are not so resistant to attack, or which use other techniques for embedding information.

- The term "fingerprint" and the phrase "embedded identifying information" describe sets of information sufficient to identify at least one designated recipient of digital content. In preferred embodiments, as described in related applications "Recovery from De-synchronization Attacks against Watermarking and Fingerprinting," (Watson) U.S. application number 10/377,266, filed 28 February 2003, and "Detecting Collusion among Multiple Recipients of Fingerprinted Information," (Watson) U.S. application number 10/378,046, filed 28 February 2003, multiple attackers colluding together cannot easily remove the fingerprint provided by the invention, or prevent at least one of them from being detected as unauthorized distributor of the digital content. However, the concept of the fingerprint as

described herein is sufficiently general to include fingerprints that are not so resistant to removal, or do not provide such capability for detecting unauthorized distributors of the digital content, or which use other techniques for embedding information, for detecting the embedded information, or for detecting unauthorized distributors of the digital content.

As described in related applications "Parallel Distribution and Fingerprinting of Digital Content," (Collens et al.) U.S. application number 10/356,692, filed 31 January 2003, and "Watermarking and Fingerprinting Digital Content Using Alternative Blocks to Embed Information," (Watson et al.) U.S. application number 10/356,322, filed 31 January 2003, a "watermark" refers to a set of locations in a media stream at which information might be embedded, while a "fingerprint" refers to the actual information that is embedded, such as for example by selecting a block or alt-block for each such location. However, in the context of the invention, there is no requirement that the concepts of watermarking and fingerprinting be so restricted. More generally, a watermark might be used for any technique by which a source of the digital content for the media stream might be identified, or a fingerprint might be used for any technique by which a recipient of the digital content for the media stream might be identified. For example, not intended to be limiting in any way, watermarking and fingerprinting information as described herein includes a representation of the entire path (or set of paths) by which the digital content representing the media stream was sent from its source and received by its end viewer (or equipment associated therewith).

- The phrase "identifying information" describes, generally, either information associated with a watermark, information associated with a fingerprint, or other information by which authorized or unauthorized distribution of digital content representing a media stream might be identified.

- The phrases "original movie" and "alt-movie" describe alternative versions of the same media stream, such as one being an original version of that media stream introduced into a system using aspects of the invention, and another being an alternative version of that same media stream generated in response to the original

movie. Similarly, the phrases "original block" and "alt-block" describe alternative versions of the same individual block or macroblock within the original movie or alt-movie. As described in related applications "Parallel Distribution and Fingerprinting of Digital Content," (Collens et al.) U.S. application number 10/356,692, filed 31 January 2003, and "Watermarking and Fingerprinting Digital Content Using Alternative Blocks to Embed Information," (Watson et al.) U.S. application number 10/356,322, filed 31 January 2003, a difference between the original movie and the alt-movie is historical, in that the alt-movie can be substituted for the original movie in nearly every respect. Similarly, a difference between any one original block and its associated alt-block is historical, in that the alt-block can be substituted for the original block in nearly every respect.

- The phrases "original digital content" and "altered digital content" (or in the latter case, "post-attack digital content") describe digital content representing media streams, in a first format (original digital content) and in a second format (altered digital content), the altered digital content having been produced in response to the original digital content and with the intent of representing substantially similar media streams, but with the effect that detecting identifying information from the original digital content is made relatively difficult. Thus, the altered digital content is a result of a de-synchronization attack on the original digital content. In preferred embodiments, the original digital content might be an actual original of some digital content before it was subject to a de-synchronization attack, or might be a constructed form of digital content, such as in response to an original movie and alt-movie, or in response to a set of original blocks and alt-blocks. For one example, not intended to be limiting in any way, the original digital content might be an average of the original movie and the alt-movie, or there might be two sets of original digital content, one for the original movie and one for the alt-movie. In one embodiment, a typical case of original digital content will include a block-by-block selection from the blocks of the original movie and the alt-movie. However, in the context of the invention, there is no particular restriction to such formats being used or included as the "original digital content" for which resynchronization is sought. Moreover, as described below, numerous variations on this theme are all within the scope and spirit

of the invention, and would be workable without undue experimentation or further invention.

- The phrase "end viewer" describes a recipient of the media stream for whom decoding of the digital content representing the media stream, and presentation of the media stream, is contemplated.

- The term "decoding" describes generating data in a form for presentation of the media stream, in response to the digital content representing the media stream in an encoded format. As described herein, the encoded format might include an industry standard encoded format such as MPEG-2. However, the concept of decoding as described herein is sufficiently general to include other encoding formats for media streams.

- The term "presentation" describes generating information in a form for viewing of the media stream, such as for example audio and visual information for viewing a movie. As described herein, presentation of a movie might include visual display of the frames or fields of motion picture, as well as audio presentation of a soundtrack associated with that motion picture. However, the concept of presentation as described herein is sufficiently general to include a wide variety of other forms of generating information for viewing.

- The term "packet" describes a portion of the digital content representing a media stream, such as for example as might be separately identifiable within that digital content 111 or transmitted when sending that digital content. In one embodiment, a "packet" indicates a contiguous sub-region of an MPEG-2 packet including picture slice data. In the context of the invention, a "packet" is not necessarily the same as an MPEG-2 packet, and a "packet" is not necessarily the same as a TCP/IP packet.

Other and further applications of the invention, including extensions of these terms and concepts, would be clear to those of ordinary skill in the art after purchasing this application. These other and further applications are part of the scope and spirit of the

invention, and would be clear to those of ordinary skill in the art without further invention or undue experimentation.

The scope and spirit of the invention is not limited to any of these definitions, or to specific examples mentioned therein, but is intended to include the most general concepts embodied by these and other terms.

System Elements

Figure 1 shows a block diagram of a system for secure presentation of media streams in response to encrypted digital content.

A system 100 includes a media stream source 110, a distribution network 120, a key server 130, and a set of customer premises equipment 140. The system 100 is disposed for presenting one more media streams, as represented by digital content associated with those media streams, to one or more particular selected users 150.

The media stream source 110 is capable of injecting a set of digital content 111, in the form of a sequence of packets 112, the sequence of packets 112 including digital content for at least one media stream intended for a user 150 of the system 100. In one embodiment, there might be more than one media stream source 110, and the media stream sources 110 are capable of injecting copies of the digital content adapted to particular selected users 150.

The distribution network 120 is disposed for transferring information between and among the media stream source 110, the key server 130, and the customer premises equipment 140. In one embodiment, the distribution network 120 includes a set of intermediate caches or sources 121, capable of receiving packets 112 from the media stream sources 110, caching or otherwise maintaining in storage information from those packets 112, and further adapting the digital content associated with those packets 112 to particular selected users 150.

Those skilled in the art will recognize, after perusal of this application, that the system 100, including the media stream source 110, the distribution network 120, and the intermediate caches or sources 121, are preferably disposed for adapting and encrypting the digital content 111 (as further described with regard to distribution of digital content representing media streams) as described related applications "Parallel Distribution and Fingerprinting of Digital Content," (Collens et al.) U.S. application number 10/356,692, filed 31 January 2003, and "Watermarking and Fingerprinting Digital Content Using Alternative Blocks to Embed Information," (Watson et al.) U.S. application number 10/356,322, filed 31 January 2003.

As further described herein, in one embodiment, not intended to be limiting in any way, the digital content 111 is encoded using an MPEG-2 encoding scheme, with selected portions of that digital content 111, representative of the media stream, encrypted as described, for example, in related applications "Parallel Distribution and Fingerprinting of Digital Content," (Collens et al.) U.S. application number 10/356,692, filed 31 January 2003, and "Watermarking and Fingerprinting Digital Content Using Alternative Blocks to Embed Information," (Watson et al.) U.S. application number 10/356,322, filed 31 January 2003. The selected portions of that digital content 111 preferably include only the portions of the digital content 111 representative of the presentable or displayable portions of the media stream, and preferably do not include any formatting data, metadata, or other descriptive data relating to the media stream, even if embedded in the encoded digital content 111 representative of that media stream.

As further described herein, in one embodiment, not intended to be limiting in any way, those portions of the digital content 111 are encoded with the effect that the sequence of packets 112 is substantially unchanged from an alternative sequence of packets 112 that might have been generated for the digital content 111, had that digital content 111 not been encrypted for distribution to the user 150. For example, this has the effect that the length of each packet 112 in the sequence of packets 112 is substantially unchanged from an alternative sequence of packets 112 that might have been generated for the digital content 111 had that digital content 111 not been encrypted for distribution to the user 150. This has the effect that the amount of intermediate state maintained for decoding that sequence of packets 112, and thus for decoding that digital content 111, is substantially unchanged from

an alternative sequence of packets 112 that might have been generated for the digital content 111, had that digital content 111 not been encrypted for distribution to the user 150.

As further described herein, in one embodiment, not intended to be limiting in any way, those portions of the digital content 111 are encoded with the effect that synchronization of audio with video within the digital content 111 is substantially unchanged from an alternative operation of synchronization of audio with video within the digital content 111 that might have been performed for that digital content 111, had that digital content 111 not been encrypted for distribution to the user 150. This has the effect that the degree of effort involved in decoding that digital content 111, any decoding steps involving synchronization of audio with video, are relatively equivalent to the degree of effort involved in an operation of synchronization of audio with video within the digital content 111 that might have been generated for the digital content 111, had that digital content 111 not been encrypted for distribution to the user 150.

As further described herein, in one embodiment, not intended to be limiting in any way, those portions of the digital content 111 are encoded with the effect that locating (or "seeking to") a selected position in a position in the media stream represented by the digital content 111 is substantially unchanged from an alternative operation of locating (or "seeking to") a selected position in a position in the media stream represented by the digital content 111 that might have been performed for that digital content 111, had that digital content 111 not been encrypted for distribution to the user 150. This has the effect that the degree of effort involved in an operation of locating (or "seeking to") a selected position in a position in the media stream represented by the digital content 111 is substantially unchanged from an alternative operation of locating (or "seeking to") a selected position in a position in the media stream represented by the digital content 111 that might have been performed for that digital content 111, had that digital content 111 not been encrypted for distribution to the user 150.

Moreover, as further described herein, in one embodiment, not intended to be limiting in any way, in the context of the invention, it is not necessary to decrypt portions of the digital content 111 to perform an operation of locating (or "seeking to") a selected position in a position in the media stream represented by the digital content 111. After

reading this application, those skilled in the art would recognize that the operation of locating (or "seeking to") a selected position in a position in the media stream represented by the digital content 111 might thus be performed relatively more efficiently (that is, without substantial additional encryption steps) and relatively more securely (that is, by relatively less trusted hardware or software components). In one embodiment, those portions of the digital content 111, in an MPEG-2 encoding of that digital content 111, useful for that operation of locating (or "seeking to") a selected position in a position in the media stream are not encrypted.

As further described herein, in one embodiment, not intended to be limiting in any way, within the digital content 111, only the video block data is encrypted, preferably using a block-substitution cipher, preferably a variation of the AES cipher, such as for example AES-128 or AES-256. In one embodiment, the block-substitution cipher can be used to encrypt each sequence of 16 bytes of video block data in each packet 112, with the fact that as many as 15 bytes of video block data within each packet 112 might remain in the clear after encryption.

In one embodiment, the digital content 111 is encoded using MPEG-2, which includes its audio and video data (as well as control data) within an MPEG "packet." MPEG packets are enclosed by MPEG-2 within an MPEG "pack." The MPEG standard is further described in documents known in the digital video industry. This has the effect that, in such embodiments, only audio or video data is encrypted (but not necessarily all audio and video data is encrypted), while substantially all of the MPEG control data (including MPEG packet headers, MPEG pack headers, and in general all types of MPEG control data), is left unencrypted. This also has the effect that, in such embodiments, only MPEG packet payloads are encrypted.

In such embodiments, where an MPEG packet includes a payload that is not an integer multiple of the encryption size (16 bytes), any remainder, possibly as many as 15 bytes, is also left unencrypted. This has the effect that, in such embodiments, at least some packets 112 might include packet header information (unencrypted), MPEG control data (unencrypted), audio or video data that is encrypted, and possibly as many as 15 bytes of audio or video data that is left unencrypted.

In such embodiments, where the MPEG data has already been encrypted with another technique (such as for example CSS, which might be in use for selected DVD physical media carrying the MPEG data), those packets 112 already encrypted with the other technique are not further encrypted using the AES cipher. Those skilled in the art will recognize that because the CSS specification provides that no more than 50% of sectors of a DVD video disk are encrypted using CSS, this has the effect that as many as 50% of sectors of the DVD video disk would remain to be possibly encrypted using the AES cipher.

In such embodiments, those data elements of the MPEG packet that have been encrypted are maintained as offsets into the MPEG pack information and MPEG packet information. This has the effect that, although the MPEG pack information and MPEG packet information have variable-length headers, the encrypted data elements can still be located relative to the end of those headers.

As further described herein, in one embodiment, not intended to be limiting in any way, within the digital content 111, separable media streams, such as for example an audio stream distinguishable from the video stream, are preferably separately encrypted, with the effect that the separable media streams might be made differently available to distinct particular selected users 150, or distinct groups of particular selected users 150.

The key server 130 is capable of supplying, such as for example in response to a request from the user 150, digital information including decryption keys (whether symmetric keys, or asymmetric keys such as used in public key cryptosystems) and license information to particular selected users 150.

The customer premises equipment 140 includes a local library 141, a local area network 142, and a set of player equipment 143. The customer premises equipment 140 is disposed for presenting one or more media streams, as represented by digital content included in the sequence of packets 112, to one or more particular selected users 150 associated with the particular selected customer premises equipment 140.

The local library 141 includes a processor 141a, program and data memory or mass storage 141b, and a formatted-media reader 141c. In one embodiment, the local library

141 also includes at least one input element 141d and at least one output element 141e. The memory or mass storage 141b is capable of including instructions 141f capable of being executed or interpreted by the processor 141a to perform steps as described herein. The memory or mass storage 141b is also capable of maintaining copies of at least portions of the digital content 111, possibly watermarked or fingerprinted as described in related applications "Parallel Distribution and Fingerprinting of Digital Content," (Collens et al.) U.S. application number 10/356,692, filed 31 January 2003, and "Watermarking and Fingerprinting Digital Content Using Alternative Blocks to Embed Information," (Watson et al.) U.S. application number 10/356,322, filed 31 January 2003.

As described below, the instructions 141f direct the local library 141 to perform the following actions:

(A1a) to receive digital content 111 from the media stream source 110, using the format of the sequence of packets 112, or

(A1b) to receive digital content 111 from the formatted-media reader 141c;

In the event that the digital content 111 is received from the formatted-media reader 141c, that digital content 111 might either be (1) already encrypted on the physical media being read by the device, (2) unencrypted on the physical media being read by the device, or (3) encrypted on the physical media being read by the device, but using a non-preferred encryption technique. In case 2, the digital content 111 is encrypted by the formatted-media reader 141c, or by an supplemental device coupled thereto, before transferring any digital content 111 to devices other than the formatted-media reader 141c. In case 3, the digital content 111 is decrypted using the non-preferred encryption technique, and re-encrypted using a preferred encryption technique, before transferring any digital content 111 to devices other than the formatted-media reader 141c.

(A2) (optionally) to partially decode that digital content 111, with the effect of retrieving at least some metadata regarding that digital content 111 in the clear, such as for example index files including pointers into the digital content 111;

(A3) to maintain that encrypted digital content 111, and optionally at least some decrypted metadata regarding that digital content 111, in the memory or mass storage 141b; and

(A4) to decode that digital content 111, with the effect of retrieving metadata regarding that digital content 111 in the clear, and with the effect of retrieving data representing presentable portions of the media stream represented by that digital content 111 in an encrypted form;

(A5) to transfer that encrypted digital content 111 from the memory or mass storage 141b to the local network 142 and to the player equipment 143; and

(A6) to decrypt selected portions of that digital content 111, in response to requests from the player equipment 143, with the effect of retrieving, in the clear but secure from detection or intrusion, data represented by that digital content 111 for presenting a media stream at the player equipment 143.

The specific techniques to be applied are further described below.

As described below, the player equipment 143 performs the following actions:

(B1) receives the decoded digital content 111 from the memory or mass storage 141b and the local network 142;

(B2) receives a set of commands or requests from the user 150;

(B3) performs those commands or requests from the user 150 capable of being performed without reference to encrypted elements of the decoded digital content 111, without performing any decryption on that decoded digital content 111; and

(B4) presents or displays those elements of the decoded digital content 111 that involve decrypting elements (such as audio or video blocks) of that decoded digital content 111, using one or more decryption keys from the key server 130.

The specific techniques to be applied are further described below.

Method of Operation

Figure 2 shows a process flow diagram of a method for secure presentation of media streams in response to encrypted digital content.

Although described serially, the flow points and method steps of the method 200 can be performed by separate elements in conjunction or in parallel, whether asynchronously or synchronously, in a pipelined manner, or otherwise. In the context of the invention, there is no particular requirement that the method must be performed in the same order in which this description lists flow points or method steps, except where explicitly so stated.

At a flow point 210, the local library 141 is ready to receive digital content 111 representing one or more media streams. The method 200 performs either the step 211 (receiving digital content 111 from the media stream source 110), or the step 212 (receiving digital content 111 from the formatted-media reader 141c).

At a step 211, the local library 141 receives digital content 111 representing one or more media streams from the media stream source 110. As part of this step, the local library 141 receives a sequence of one or more packets 112, collectively including the digital content 111. As part of this step, the local library 141 might be required to request retransmission of lost or broken packets 112, might be required to reorder packets 112 delivered out of sequence, and might be required to re-establish a connection with the media stream source 110 to continue receiving from a known breakpoint. As a result of this step, the local library 141 obtains at least a portion of the digital content 111 representing one or more media streams, and the method 200 is able to proceed at the flow point 220.

At a step 212, the local library 141 receives digital content 111 representing one or more media streams from the formatted-media reader 141c. As part of this step, the local library 141 receives data directly from the formatted-media reader 141c or from a supplemental device coupled thereto. That data might be delivered in a sequence of one or more packets 112, in a similar manner to performance of the step 211, or might be delivered by another technique, such as for example a DMA transfer. As noted above, that digital content 111 might either be already encrypted, unencrypted, or encrypted using a non-preferred encryption technique. As part of this step, as noted above, the digital content 111 is ultimately transformed into a format using a preferred encryption technique before being transferred to any devices other than the formatted-media reader 141c. As a result of this step, the local library 141 obtains at least a portion of the digital content 111 representing one or more media streams, and the method 200 is able to proceed at the flow point 220.

At a flow point 220, the local library 141 is ready to partially decode the digital content 111. Steps following this flow point are optionally performed as part of the method 200.

At a step 221, the local library 141 partially decodes the received digital content 111, with the effect of obtaining, in the clear, at least some metadata regarding that digital content 111. In one embodiment, the metadata obtained in the clear includes at least one index file including pointers to selected locations within the media stream represented by the digital content 111. The method 200 is able to proceed at the flow point 230.

At a flow point 230, the local library 141 is ready to maintain digital content 111 in the memory or mass storage 141b.

At a step 231, the local library 141 records the digital content 111 in the memory or mass storage 141b.

At a step 232 (if the steps following the flow point 220 were performed), the local library 141 records any metadata obtained in response to the digital content 111 in the memory or mass storage 141b.

As a result of performing the steps following the flow point 230, the local library 141 is able to retrieve the encrypted digital content 111, and optionally at least some unencrypted metadata associated therewith, from the memory or mass storage 141b. The method 200 is able to proceed with the flow point 240.

At a flow point 240, the local library 141 is ready to send the encrypted digital content 111 to the player equipment 143.

At a step 241, the local library 141 retrieves the encrypted digital content 111, and optionally at least some unencrypted metadata associated therewith, from the memory or mass storage 141b.

At a step 242, the local library 141 sends that encrypted digital content 111 from the memory or mass storage 141b, using the local network 142, to the player equipment 143.

As a result of performing the steps following the flow point 240, the player equipment 143 is able to access the encrypted digital content 111. The method 200 is able to proceed with the flow point 250.

At a flow point 250, the player equipment 143 is ready to present the encrypted digital content 111 to the user 150.

At a step 251, the player equipment 143 receives the encrypted digital content 111, using the local network 142, from the memory or mass storage 141b.

At a step 252, the player equipment 143 receives a set of commands or requests from the user 150.

At a step 253, the player equipment 143 performs those commands or requests from the user 150 capable of being performed without reference to encrypted elements of the decoded digital content 111, without performing any decryption on that

decoded digital content 111. As part of this step, the player equipment 143 might perform one or more of the following sub-steps:

At a sub-step 253a, the player equipment 143 might rewind, fast forward, or otherwise "seek" to a selected location within the digital content 111.

At a sub-step 253b, the player equipment 143 might pause or halt presentation of the media stream represented by the digital content 111.

At a step 254, the player equipment 143 performs those commands or requests from the user 150 to perform the media stream represented by the digital content 111. To perform this step, the player equipment 143 performs the following sub-steps:

At a sub-step 254a, the player equipment 143 decodes the digital content 111, with the effect of obtaining metadata describing presentation of the media stream, and encrypted data for presentation of the actual audio and video associated with the media stream.

At a sub-step 254b, the player equipment 143 sends encrypted digital content 111 to a supplemental device (or a secure sub-system) for decryption.

At a sub-step 254c, the player equipment 143 receives decrypted digital content 111 from the supplemental device (or the secure sub-system) after decryption.

At a sub-step 254d, the player equipment 143 presents the media stream in response to the decrypted digital content 111.

At a flow point 260, the player equipment 143 is ready to respond to further commands from the user 150, and is able to proceed with the flow point 250.

Alternative Embodiments

The invention is useful for, and has sufficient generality for, applications other than distribution of media streams, and to other than distribution of digital content. For example, the invention is also generally useful for applications in which security of datasets or identifying recipients of those datasets is desired.

Although preferred embodiments are disclosed herein, many variations are possible which remain within the concept, scope, and spirit of the invention. These variations would become clear to those skilled in the art after perusal of this application.

- As noted above, the invention is not restricted to movies, but is also applicable to other media streams, such as for example animation or sound, as well as to still media, such as for example pictures or illustrations, and to databases and other collections of information.

Those skilled in the art will recognize, after perusal of this application, that these alternative embodiments are illustrative and in no way limiting.

Claims

1. A method, including steps of

encoding a media stream into a digital content format representing that media stream; and

encrypting a portion of that digital content, less than the entire digital content format representing that media stream, the portion of the digital content that is encrypted being required for presentation of the media stream;

wherein the encrypted version of that digital content is substantially unchanged in formatting parameters from an unencrypted version of that digital content.

2. A method as in claim 1, wherein

said steps of encoding provide an MPEG encoding of at least some video data.

3. A method as in claim 1, wherein

said steps of encrypting include steps of

encrypting at least some audio or video data using a block-substitution cipher.

4. A method as in claim 1, wherein

said steps of encrypting include steps of

encrypting at least some audio or video data using a block-substitution cipher; and

refraining from encrypting at least some audio or video data using that block-substitution cipher, wherein an amount of audio or video data not encrypted is less than a block size for that block-substitution cipher.

5. A method as in claim 1, wherein

said steps of encrypting include steps of

identifying at least a first set of data and a second set of data in the digital format; and

separately encrypting the first set of data and the second set of data;

whereby the first set of data can be made available to a first set of users and the second set of data can be made available to a second set of users, the first set of users being distinguishable from the second set of users.

6. A method as in claim 1, wherein

said steps of encrypting include steps of

refraining from encrypting at least one of (a) information by which at least some audio or video data is described, or (b) at least some formatting information.

7. A method as in claim 1, wherein

the digital content format includes

at least some audio or video data; and

at least some formatting information.

8. A method as in claim 1, wherein

the digital content format representing that media stream includes a set of layers, each relatively higher-level layer representing an abstraction for which each relatively lower-level layer represents an implementation thereof;

a first set of relatively higher-level layers represent audio or video information for the media stream, while a second set of relatively lower-level layers represent techniques by which that information is formatted or supplemented; and

the step of encrypting is applied only to that portion of the digital content representing audio and video information.

9. A method as in claim 1, wherein

the digital content format representing that media stream includes a set of layers, each relatively higher-level layer representing an abstraction for which each relatively lower-level layer represents an implementation thereof;

a first set of relatively higher-level layers represent audio or video information for the media stream, while a second set of relatively lower-level layers represent techniques by which that information is broken into packets, indexed, multiplexed, or supplemented with metadata; and

the step of encrypting is applied only to that portion of the digital content representing audio and video information.

10. A method as in claim 1, wherein

the digital content format representing that media stream includes a set of layers, each relatively higher-level layer representing an abstraction for which each relatively lower-level layer represents an implementation thereof;

a first set of relatively higher-level layers represent audio and video information for the media stream, while a second set of relatively lower-level layers represent techniques by which that information is broken into packets, indexed, multiplexed, or supplemented with metadata; and

the step of encrypting is not applied to at least part of that portion of the digital content representing other than audio and video information.

11. A method as in claim 1, wherein

the media stream includes at least one of: a movie, animation, sound, still media, a picture, an illustration, a database, a collection of information.

12. A method as in claim 1, including steps of

selecting that portion of the digital content for encryption so there is no substantial change in distribution of that digital content.

13. A method as in claim 12, wherein

said steps of selecting include ensuring there is no substantial change in packetization of a set of digital data in that digital content.

14. A method as in claim 12, wherein

said steps of selecting include ensuring there is no substantial change in synchronization of audio with video portions of the media stream.

15. A method as in claim 12, wherein

said steps of selecting include ensuring there is no substantial change in length of at least some identifiable audio or video data in that digital content.

16. Apparatus including

an input port capable of being coupled to a communication link, the communication link being capable of carrying digital content, the digital content including at least some presentable information and at least some formatting information;

a digital content decoder, the decoder being capable of identifying the presentable information in response to the formatting information;

a digital content decryptor, the decryptor being capable of decrypting the presentable information in response to a key;

wherein the decryptor is protected by a relatively-higher degree of security than the decoder.

17. Apparatus as in claim 16, wherein the communication link includes at least one of:

a computer network capable of carrying digital content;

a reader capable of retrieving information in response to physical media, the physical media being capable of carrying digital content.

18. Apparatus as in claim 16, wherein the decoder includes an MPEG decoder.

19. Apparatus as in claim 16, wherein

the decoder is included in a first selected set of hardware or software, the first selected set being trusted; and

the decryptor and the key are included in a second selected set of hardware or software, the second selected set being relatively more trusted than the first selected set.

20. Apparatus as in claim 16, wherein the decoder is responsive to the formatting information to present at least some metadata about one or more media streams without the decoder having access to the presentation information.

21. Apparatus as in claim 16, wherein the decoder is responsive to the formatting information to provide at least one of the following functions without the decoder having access to the presentation information:

known playback functions known for media streams;

navigation within the digital content;

content selection within the digital content; or

manipulation of the presentation.

22. Apparatus as in claim 16, wherein the digital content represents a media stream including at least one of: a movie, animation, sound, still media, a picture, an illustration, a database, a collection of information.

23. Apparatus as in claim 16, wherein the relatively-higher degree of security includes tamper-resistant hardware operating under control of verified software.

24. Apparatus as in claim 16, wherein

the digital content represents a first media stream and a second media stream,

the decoder being responsive to the formatting information and the decryptor being responsive to a selected key,

the selected key providing differential access to selected users to the first media stream and the second media stream.

25. Apparatus as in claim 24, wherein

the first media stream includes audio information and the second media stream includes video information;

the first media stream includes information in a first language and the second media stream includes information in a second language;

the first media stream includes presentation information targeted at a first type of audience and the second media stream includes information targeted at a first type of audience;

26. A method, including steps of

encoding a media stream into a digital content format representing that media stream, that digital content format having a set of information nodes, those information nodes being disposed in at least a partial ordering;

encrypting a portion of that digital content, the portion being encrypted less than the entire digital content format representing that media stream, the portion of the digital content that is encrypted being required for presentation of the media stream;

wherein the unencrypted portion of that digital content is substantially closed in a direction under that partial ordering, whereby it is possible to decode the unencrypted portion of that digital content without having to decrypt it.

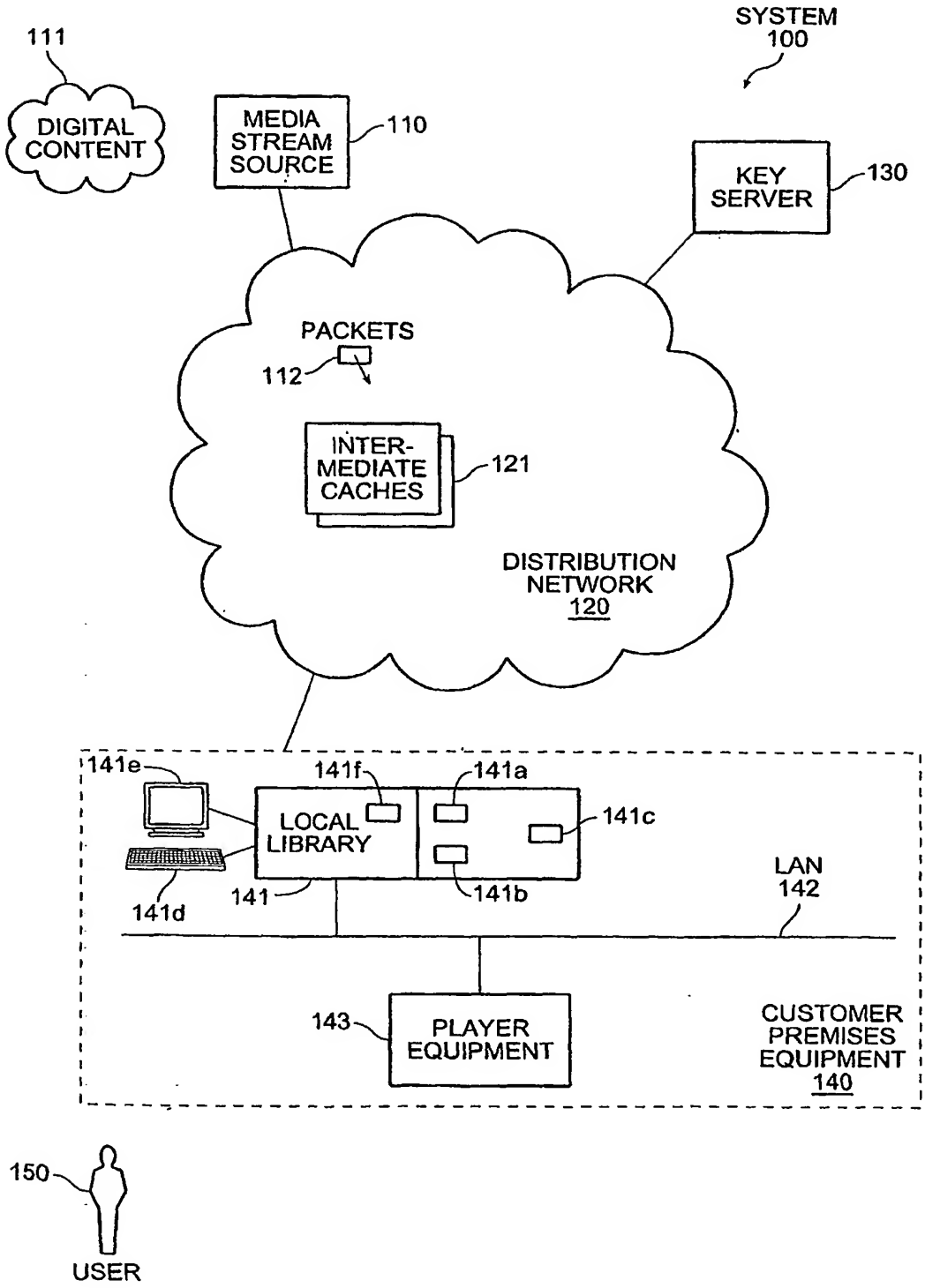

111
DIGITAL CONTENT
SYSTEM 100
MEDIA STREAM SOURCE
110
KEY SERVER
130
PACKETS
112
INTER-MEDIATE CACHES
121
DISTRIBUTION NETWORK 120
141e
141f
141a
LOCAL LIBRARY
141c
141d
141
141b
LAN 142
143
PLAYER EQUIPMENT
CUSTOMER PREMISES EQUIPMENT 140
150
USER

FIG. 1

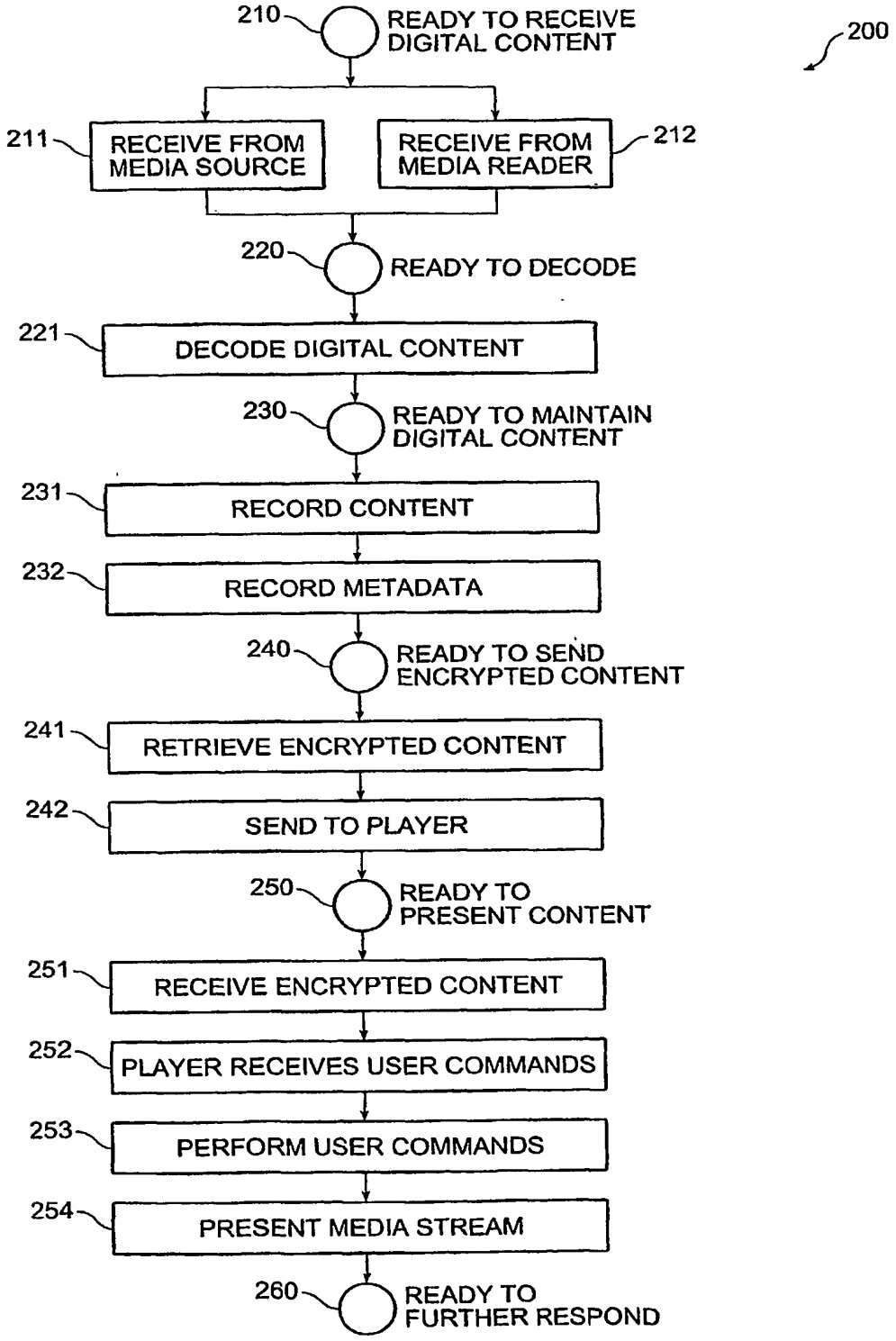

200
210
READY TO RECEIVE DIGITAL CONTENT
211
RECEIVE FROM MEDIA SOURCE
212
RECEIVE FROM MEDIA READER
220
READY TO DECODE
221
DECODE DIGITAL CONTENT
230
READY TO MAINTAIN DIGITAL CONTENT
231
RECORD CONTENT
232
RECORD METADATA
240
READY TO SEND ENCRYPTED CONTENT
241
RETRIEVE ENCRYPTED CONTENT
242
SEND TO PLAYER
250
READY TO PRESENT CONTENT
251
RECEIVE ENCRYPTED CONTENT
252
PLAYER RECEIVES USER COMMANDS
253
PERFORM USER COMMANDS
254
PRESENT MEDIA STREAM
260
READY TO FURTHER RESPOND

FIG. 2